

# Contents

<b>1</b>	<b>Econometrics</b>	<b>3</b>
1.1	Low and Meghir, 2017, JEP . . . . .	3
1.2	Lewbel, 2016, JEL . . . . .	4
1.2.1	Point Identification . . . . .	4
<b>2</b>	<b>IO</b>	<b>5</b>
2.1	Rey and Stiglitz, 1995, RAND . . . . .	5
2.1.1	Benchmark Case . . . . .	6
2.2	Production Function Estimation . . . . .	7
<b>3</b>	<b>Decision Theory</b>	<b>9</b>
3.1	Dekel and Lipman, 2010, Annu Rev Econ . . . . .	9
<b>4</b>	<b>Matching Theory</b>	<b>11</b>
4.1	Gale and Shapley, 1962 . . . . .	11
<b>5</b>	<b>Macro</b>	<b>13</b>
5.1	Kiyotaki, 2011 . . . . .	13



# Chapter 1

## Econometrics

### 1.1 Low and Meghir, 2017, JEP

Hamish Low and Costas Meghir. The Use of Structural Models in Econometrics. *Journal of Economic Perspectives*, 31(2):33–58, May 2017

#### Defining a structural model:

A **fully specified model** make explicit assumptions about the economic actors' objectives and their economic environment and information set, as well as specifying which choices are being made within the model. They allow a complete solution to the individual's optimization problem as a function of current information set. Fully specified models are particularly useful in understanding mechanism of a policy, especially when we want to estimate some long-term effects of the policy.

A **partially specified model** relies on a sufficient statistic that summarizes choices not being modeled specifically. For example, assuming that the choices is only intratemporal instead of intertemporal.

**Treatment effect models** focus on identifying a specific causal effect of a policy while saying least about the theoretical environment. The pro is the cleanness of causality. The con is the limitation in exploiting the results outside. The identification of treatment model depends on assumptions that the experientment has not been compromised and there is no spillovers from the treatment units.

A combination of fully specified model and randomized experiments can enhance analysis for both. Experimental evidence can be used either to validate a structural model, or to aid in the estimation process (in identification).

#### Solving structural models

This has been described in Adda and Cooper(2003) well. The general process is: 1) write down the bellmand function; 2) discrete the state space and decision space; 3) use value function iteration to solve the bellman function.

## 1.2 Lewbel, 2016, JEL

A Lewbel. The Identification Zoo-Meanings of Identification in Econometrics. *Forthcoming on Journal of Economic Literature*, 2016

<https://www2.bc.edu/arthur-lewbel/ident-zoo-SL-Part1.pdf>

<https://www2.bc.edu/arthur-lewbel/ident-zoo-SL-Part2.pdf>

are links for a ppt on this paper by Lewbel.

There are two kinds of identification problems. 1. One is to identify the treatment effect, a typical example is the selection bias. The problem in these cases are that selection (determining who is treated or observed) and outcomes may be correlated. 2. Another is to identify the true coefficient in a linear regression when regressors are measured with error.

### 1.2.1 Point Identification

We start by assuming some information  $\phi$  is knowable. A simple definition of point identification is that a parameter  $\theta$  is point identified if, given the model, is uniquely determined from  $\phi$ . Notice that this definition of point identification is recursive in some sense. To identify  $\theta$ , we first need to assume some  $\phi$  is knowable, which means  $\phi$  itself is identified. This identification of  $\phi$  can only be justified by further assumptions of DGP (Data Generating Process).

For example, for a model  $Y = X\theta + e$ , we assume that  $E(e^2) \neq 0$  and  $E(eX) = 0$ , and suppose  $\phi$  includes the second order of  $(Y, X)$ . Then we can conclude that  $\phi$  is point identified, given by  $E(XY)/E(X^2)$ . Notice that the identification comes from the *assumptions* of model.

One common DGP is IID. Under this DGP, we can consistently identify the distribution of observation  $W$ . Another DGP is where each data point consists of a value of  $X$  chosen from its support, then we randomly draw  $Y$  conditional on  $X$ , which is independent from other draws conditional on this  $X$ . Under this DGP, we can consistently identify  $F(Y|X)$ . We can also use more complicated DGPs, for example, we generally assume only the second order moments are knowable in time series. One reason is this being sufficient for identification, another reason is higher order moments become unstable over time. Assumptions over GDP are always needed, even in experient data, and which specific assumptions to take depend on the model.

# Chapter 2

## IO

### 2.1 Rey and Stiglitz, 1995, RAND

**Patrick Rey and Joseph Stiglitz. The Role of Exclusive Territories in Producers' Competition. *The RAND Journal of Economics*, 26(3):431–451, October 1995**

For detailed proof of this paper, see Evernote.

**Main result:** vertical restrains can be used to reduce interband competition. Because exclusive territories alter the perceived demand curve, making each producer believe he faces a less elastic demand curve, inducing an increase in eqm price and producer's profits even in the absence of franchise fee. This result is different from traditional Chicago school results, which insist that exclusive territories will increase efficiency. This difference comes from market structure. Chicago schools investigate in full competition and full monopoly producer cases, while this paper looks at duopoly producer. In full competition and full monopoly case, the competition level has already been *preassumed*, while exclusive territories can reduce the competition level among producers in other cases.

The key for the result is the the following compound demand elastic:

$$\tilde{\varepsilon}(p^e) := m_1(p, p)\varepsilon_1(q, q) + m_2(p, p)\varepsilon_2(q, q)$$

where  $q = q_1^r(p, p)$  (the response retail price),  $m_1(p, p) = \partial \log q_1^r(p_1, p_2) / \partial \log p_1$  (the own elasticity of retail price to producer price), and  $m_2(p, p) = \partial \log q_1^r(p_1, p_2) / \partial \log p_2$  (the cross elasticity of retail price to producer price), and  $\varepsilon_1(q_1, q_2) = -\partial \log D^1(q_1, q_2) / q_1$  (the own elasticity of demand to retail price, positive), and  $\varepsilon_2(q_1, q_2) = -\partial \log D^1(q_1, q_2) / q_2$  (the cross elasticity of demand to retail price, negative).

In the above equation, it is very reasonable to think  $0 < m_1 < 1$  and  $0 < m_2$ .  $m_1 > 0$  because the own elasticity of retail price to producer price is positive.  $m_1 < 1$  means that retailer will absorb some increase in producers' price, which will be the case if demand elasticity becomes higher in high retail price. And

$0 < m_2$  derives from the two products to be substitutes. Under these two conditions, combined with  $\varepsilon_1 > 0$  and  $\varepsilon_2 < 0$ , then  $\tilde{\varepsilon}(p^e) < \varepsilon_1$ . Thus, under exclusive territories, the producers' perceived demand curve is less elastic. So the equilibrium price (both retail and producer) is higher even when no franchise fee applies.

**Setting:**

- two manufacturers produce imperfect substitutes at same marginal cost  $c$
- retailers are perfect competition / or exclusive territory
- the final good demand depends on retail prices and is given by  $D^i(q_1, q_2)$
- costs and demand functions are common knowledge, retailers observe all contracts signed by each producer
- producers only observe the quantity bought by retailers; they do not observe the quantities sold by retailers (i.e. fullline forcing is infeasible)
- producers serve many markets at no additional cost
- consumers have no search cost

Under these settings and information conditions, we can see it as a two stage game. In the first stage, producers simultaneously set wholesale price  $p_1$  and  $p_2$ . In the second stage, the retailer observe all wholesale price and decide the retail price simultaneously.

### 2.1.1 Benchmark Case

We use the following assumptions throughout the paper unless specified otherwise:

1. Let  $\pi(p_i, q_1, q_2) := (q_i - p_i)D^i(q_1, q_2)$  denote the retail profit for product  $i$ ; assume it to be twice differentiable wrt each argument, and is single peaked wrt  $q_i$ . The reaction function  $q_i^a(p_i, q_j)$  is thus continuously differentiable and characterized by FOC.
2. Products are substitutes:  $\partial D^i / \partial q_i \leq 0$  and  $\partial D^i / \partial q_j \geq 0$
3. Demand functions are symmetric:  $\forall p_1, p_2 \in \mathbb{R}_+, D^1(p_1, p_2) = D^2(p_2, p_1)$

Think of the benchmark case that no vertical restriction so the retailers are perfect competitive, and the producer monopolizes. In this case, the game is just a one step optimal pricing problem. The producer chooses an optimal retail price.

**Useful trick:**

Throughout this paper, we can write the symmetric eqm conditions in the following form:

$$(p^c - c)/p^c = 1/\varepsilon(p^c, p^c)$$

where  $p^c$  is the symmetric eqm price, and  $\varepsilon$  is some kind of elasticity.

In a symmetric eqm, the FOC of each producer gives  $(p^c - c)/p^c = 1/\varepsilon_1(p^c, p^c)$  where  $\varepsilon_1 = -\partial \log D^1(q_1, q_2)/\partial \log q_1$ , i.e. the self elasticity of demand.

In the simplest benchmark case, the two factories are integrated, leading us to:  $(p^c - c)/p^c = 1/E(q^m)$ ,  $E(q) := \varepsilon_1(q, q) + \varepsilon_2(q, q)$ , where  $\varepsilon_2 = -\partial \log D^1(q_1, q_2)/\partial \log q_2$  (the cross demand elasticity).

In the exclusive territory case, the symmetric eqm satisfies  $(p^e - c)/p^e = 1/\tilde{\varepsilon}(p^e)$ , where

$$\tilde{\varepsilon}(p^e) := m_1(p, p)\varepsilon_1(q, q) + m_2(p, p)\varepsilon_2(q, q)$$

where  $q = q_1^r(p, p)$  (the response retail price),  $m_1(p, p) = \partial \log q_1^r(p_1, p_2)/\partial \log p_1$  (the own elasticity of retail price to producer price), and  $m_2(p, p) = \partial \log q_1^r(p_1, p_2)/\partial \log p_2$  (the cross elasticity of retail price to producer price).

## 2.2 Production Function Estimation

There are several difficulties in estimating production functions which will cause endogeneity problem.

To name some:

- Simultaneouty. For a production function like  $y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 l_{it} + \omega_{it} + \varepsilon_{it}$ , we suppose that  $\omega_{it}$  is obseved by firm but not by econometricians, and  $\varepsilon_{it}$  is observed neither by firm nor econometrician. This could induce endogeneity when directly using OLS. Suppose  $\omega_{it}$  has positive serial correlation, when  $\omega_{it}$  is high, firm will choose higher input bacause it is very possible that  $\omega$  in next period will be high as well. This induces a positive bias for  $\beta_k$ .
- Exit selection bias. When constructing a balanced panel, then we only observe those firms with high  $\omega_{it}$ .

There are several approaches in solving this problem.

- Classical Approach.
  - Use input price as IV. The problem is input price is generally the same across firms. When we do observe different input prices, which generally indicates market power in input markets. Thus those facing higher  $\omega_i$  will tend to produce more, inducing their input price to increase, which has a negative bias on  $\beta_k$ .
  - Use fixed effect. If we assume that a firm faces same  $\omega_{it}$  across time, which is not very practical.

- Proxy for  $\omega_{it}$ . This approach first raised by Olley and Pakes (1996) and developed by Levinsohn and Petrin (2003) and Akerberg et al. (2015). We name them OP, LP, ACF here after. The key idea is to use a function of what we know (for example,  $k_t$  and  $i_{t-1}$  in OP) to proxy  $\omega$ , the existence of such function is guaranteed by the solution property of the firm optimization problem. And the specific form of this function is estimated nonparametrically.



## Chapter 3

# Decision Theory

### 3.1 Dekel and Lipman, 2010, Annu Rev Econ

E Dekel and B L Lipman. How (not) to do decision theory. *Annu Rev Econ*, 2(1):257–282, 2010

#### What we can learn from decision theory

1) to know whether our intuition is correct; 2) to flesh out initial intuition to get additional or better prediction (e.g. some additional observable prediction); 3) to help us understand of why and how mechanism works

#### Story of the model

Why the story of the model is important even if it is not realistic? Why we can't just choose the model by the 'prediction-rejection' procedure?

1) The story makes us believe the predictions more; 2) Having a nice intuition helps us to utilize the model and to expand the model; 3) And a model can never be realistic, the important thing is whether it captures something important.



## Chapter 4

# Matching Theory

### 4.1 Gale and Shapley, 1962

D Gale and L S Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, January 1962

This paper introduces Deferred Acceptance (DA) algorithms in matching, and shows that it is stable and optimal.

**Definition 4.1.1.** An assignment is called **unstable** if there are two applicants  $\alpha$  and  $\beta$  who are assigned to universities A and B, although  $\beta$  prefers A to B and A prefers  $\beta$  to  $\alpha$ .

**Definition 4.1.2.** A stable assignment is called **optimal** if *every* applicant is at least as well off under it as under any other stable assignments.

Comment: it looks at first that optimal assignments may not always exist. But as we will show constructively, it does.

**Definition 4.1.3. Deferred Acceptance algorithm** works as following. First let each boy propose to his favorite girl. Each girl who receives more than one proposal rejects all but her favorite boy. Yet, she doesn't accept him, but keeps him in a string to allow for the better may come later. Then in the second round, each rejected boy propose their second favorite girl, and each girl picks her favorite among the new proposers and the one in the string, and put him in her string. This process continues, until each girl has one boy in the string (this will happen after finite rounds), then each girl picks the boy in her string.

We will show the assignment by above algorithm is stable and optimal.



## Chapter 5

# Macro

### 5.1 Kiyotaki, 2011

**N Kiyotaki. A mechanism design approach to financial frictions. *manuscript*, 2011**

This paper mainly discusses the micro structures which may induce financial friction (inefficiency). These structures include private information of income, private technology for storage, limited commitment and their combinations.



# Bibliography

- Daniel A Akerberg, Kevin Caves, and Garth Frazer. Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451, 2015.
- E Dekel and B L Lipman. How (not) to do decision theory. *Annu Rev Econ*, 2(1):257–282, 2010.
- D Gale and L S Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, January 1962.
- N Kiyotaki. A mechanism design approach to financial frictions. *manuscript*, 2011.
- James Levinsohn and Amil Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341, April 2003.
- A Lewbel. The Identification Zoo-Meanings of Identification in Econometrics. *Forthcoming on Journal of Economic Literature*, 2016.
- Hamish Low and Costas Meghir. The Use of Structural Models in Econometrics. *Journal of Economic Perspectives*, 31(2):33–58, May 2017.
- G Steven Olley and Ariel Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297, November 1996.
- Patrick Rey and Joseph Stiglitz. The Role of Exclusive Territories in Producers’ Competition. *The RAND Journal of Economics*, 26(3):431–451, October 1995.